

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 23, 2003, 17:23:34 ; Search time 2961 Seconds  
(without alignments)  
4835.653 Million cell updates/sec

Title: US-09-745-506-37  
Perfect score: 1799  
Sequence: 1 MDKALSLSLNDFASLSFAE.....LENKINIILSETDRDPIQVY 350

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODER=frame+ p2n.model -DEV=tlh  
-Q/cgn2.1/USPTO\_spool/US09745506/runatc.22082003\_104401\_7037/app-query.fasta.1.519  
-DB=genemb1 -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cd1 -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09745506.ecgn.1.1.2377/runatc.22082003\_104401\_7037 -NCPD=6 -ICPD=3  
-NO\_MMAP -LARGEDUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: gb\_ba:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_scs:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_hlg\_hum:\*  
31: em\_hlg\_inv:\*  
32: em\_hlg\_other:\*  
33: em\_hlg\_mus:\*  
34: em\_hlg\_pin:\*  
35: em\_hlg\_rod:\*  
36: em\_hlg\_mam:\*  
37: em\_hlg\_vrt:\*  
38: em\_sy:\*  
39: em\_hugo\_hum:\*  
40: em\_hugo\_mus:\*  
41: em\_hugo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1799	100.0	1053	6	AX119075	AX119075 Sequence
2	1799	100.0	1387	9	AF060513	AF060513 Homo sapi
3	1799	100.0	1425	9	AF283538	AF283538 Homo sapi
4	1799	100.0	1440	9	BC007654	BC007654 Homo sapi
5	1799	100.0	1574	6	BD158389	BD158389 Homo sapi
6	1799	100.0	1574	9	AK023378	AK023378 Homo sapi
7	1799	100.0	1606	9	AB038949	AB038949 Homo sapi
8	1763	98.0	1579	9	AF182416	AF182416 Homo sapi
9	1715	95.3	1353	9	HS0805522	AL834430 Homo sapi
10	1575	87.5	1836	10	AF284439	AF284439 Mus muscu
11	976.5	54.3	55374	9	AL645474	AL645474 Human DNA
12	826.5	45.9	796	6	BD149184	BD149184 primer fo
13	810	45.0	1328	3	AK114307	AK114307 Clona int
14	737	41.0	170588	2	AC037455	AC037455 Homo sapi
15	737	41.0	190508	9	AC005037	AC005037 Homo sapi
16	737	41.0	198250	2	AC093681	AC093681 Homo sapi
17	601	33.4	231600	2	AC130779	AC130779 Rattus no
18	585	32.5	149819	2	AC121091	AC121091 Mus muscu
19	585	32.5	234976	2	AC118698	AC118698 Mus muscu
20	578	32.1	1036	3	AY069556	AY069556 Drosophi
21	578	32.1	70667	2	AC017306	AC017306 Drosophi
22	578	32.1	165839	3	AC092398	AC092398 Drosophi
23	578	32.1	273414	3	AE003650	AE003650 Drosophi
24	578	32.1	299935	3	DROSADH09	AE003415 Drosophi
25	398	22.1	155127	2	AL929495	AL929495 Dario rer
26	386	21.5	249	6	BD049005	BD049005 Sequence
27	384.5	21.4	1187	8	SCYGL221C	272743 S. cerevisia
28	358.5	19.9	12320	1	AE013129	AE013129 Thermomara
29	353	19.6	901	8	SC052042	US2042 Saccharomyc
30	351	19.5	28324	8	SPPC126	AL034490 S. pombe c
31	317.5	17.6	300050	1	AP004599	AP004599 Oceanobac
32	312	17.3	300029	1	AE016748	AE016748 Staphyloc
33	305	17.0	3041	1	AF270328	AF270328 Staphyloc
34	305	17.0	3041	6	AX145646	AX145646 Sequence
35	303	16.8	302173	1	AE016951	AE016951 Enterococ
36	297	16.5	301550	1	AP003192	AP003192 Clostridi
37	296	16.5	1083	6	AX143427	AX143427 Sequence
38	288.5	16.0	281017	2	AC123462	AC123462 Rattus no
39	288.5	16.0	290029	1	AE017038	AE017038 Bacillus
40	288	16.0	3441	1	AF269787	AF269787 Staphyloc
41	288	16.0	3441	6	AX145085	AX145085 Sequence
42	285.5	15.9	218470	1	BSUB0013	AE011011 Bacillus su
43	284.5	15.8	302422	1	AE017011	AE017011 Bacillus su
44	281.5	15.6	301439	1	AE015943	AE015943 Clostridi
45	279	15.5	333750	1	AP004827	AP004827 Staphyloc

RESULT 1

## ALIGNMENTS



```

/db_xref="GI:12001976"
/translation="MDIKALLSLNDFASLSPAESMDNVLVEPSPRYNTVLTFLTN
DLTEFMEVLOKADILSYRPIFRPKRIKRNKRLVIFALENVGIXSPETA
YDAPOGVNMLAKGLGACTSRPIHPSKAPNPTBGNHVEVNTTOLDKYMAYK
GIDGVSVTSFSAKNEQTRINLNCOTALMOVDLSRNOLQOKTILSELKLL
LMTGRLCTLDESVSLAMIDRIKRLHSLRLAGVRLKLEISQVKAACAGSGS
SVLQGVADLYLTGEMSHHDTLDAASOGINILCEHSNTERGFI.SDLRMLDSLENK
INITSETRDLPLVY"
BASE COUNT      378 a      323 c      309 g      377 t
ORIGIN
Alignment Scores:
Pred. No.:      6.2e-146      Length:      1387
Score:          1799.00      Matches:      350
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             9      Gaps:      0
US-09-745-506-37 (1-350) x AF060513 (1-1387)

```

```

QY      1 MetAspLeuLysAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
DB      82 ATGGATTGAAAGGCTCTCTTCTCTGGAATGACTTTCATCCCTCTCTGCTGAG 141
QY      21 SerTPAspAsnValGlyLeuLeuValGluProSerProProHisThrValAsnThrLeu 40
DB      142 ACTTGGGACAAATGTTGGATTACTGCTGGAAACCAAGCCCACTCATCTGTAATACACTC 201
QY      41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGlnLysAlaAsp 60
DB      202 TTCTGACCAATGACTGACTGAGGAAGATGATGAGAGAGTGTGCAAAAAGAGGACAGAC 261
QY      61 LeuIleLeuSerTyrHisProProIlePheArgProMetLysArgIleThrTrpAsnThr 80
DB      262 CTCATTCTCTCTACCATCGGCTATCTCCGACCATGAAAGGCAATACCTGGAACACA 321
QY      81 TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis 100
DB      322 TGGAAAGACGCGCTGGTATCCGGGCTCTGGAGAACAGAGTGGTATCTACTCTCTCAT 381
QY      101 ThrAlaTyrAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeuGlyAla 120
DB      382 ACAGCCTATGATGCTCGCCCGAGGCGTCACAACTGGTGGCTAAAGGCTTGAGGCT 441
QY      121 CysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGlyAsnHis 140
DB      442 TGTACTCTCAGGCGCCATCATCTCTCAAAAGCTCCCACTACCTCAGAGGGAACACAC 501
QY      141 ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaValLys 160
DB      502 CGAGTGAATTCACACTTAACACCAACCAAGACTGGACAAAGTCACTGCTCAGAGTAA 561
QY      161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGluGlnThr 180
DB      562 GGAATTTGACGGTCTCTCTCACTCTTTCTGCTAGACACTGGTAATAGGAACAACA 621
QY      181 ArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGlnValValAspPheLeuSerArg 200
DB      622 CGGATTAATCTGATGTACTCAGAAAGGCTTGATCAGAGTGTAGATTTCTTCCCG 681
QY      201 AsnLysGlnLeuTyrGlnLysThrGluIleLeuSerLeuGluLysProLeuLeuLeuHis 220
DB      682 AACAAACACACTTATCAGAAAGGAAATTCGTCTACTGAGAGAGCTTGTCTTCTACAT 741
QY      221 ThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetIleAsp 240
DB      742 ACTGGGATGGGAGCGGTATGACACACTGGAATGATCTGTCTCCCTGCAACCATGATTGAT 801
QY      241 ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyAlaGlyArgThr 260
DB      802 CGAATATAAAGACACCTAAACATATCTCATATTCGTGCTTGGCCCTTGGGGTGGAGAACCC 861

```

```

QY      261 LeuGluSerGlnValLysValAlaLeuCysAlaGlySerGlySerValLeuGln 280
DB      862 TTGAGATCTCAAGTCAAAGCGCGGCGCTGTGCTGCTTGGAGACAGCTTTCGACAG 921
QY      281 GlyAlaGluAlaAspLeuTyrLeuThrGlyGluMetSerHisHisAspThrLeuAspAla 300
DB      922 GGTGTGAGGCTGACCTTATCACTCAGAGGTGAGATGTCCCATCATGATATTGATGCT 981
QY      301 AlaSerGlnGlyIleAsnValIleLeuCysGluHisSerAsnThrGluArgGlyPheLeu 320
DB      982 GCTTCCCAAGGATTAATGATCATCTCTGGAACACAGCAACACTGACAGGCTTTCCT 1041
QY      321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleLeuSer 340
DB      1042 TCTGACCTTCGACATATGCTGATCTCTCACTTGGAGAAATAGATTAATATCTATCA 1101
QY      341 GluThrAspArgAspProLeuGlnValVal 350
DB      1102 GAGACTGACAGGAGGACCTCTTCAGGTGCTA 1131

```

```

RESULT 3
AF283538
LOCUS      AF283538      1425 bp      mRNA      linear      PRI 16-JAN-2001
DEFINITION Homo sapiens NIF3L1 protein mRNA, complete cds.
ACCESSION AF283538
VERSION AF283538.1 GI:12006402
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 1425)
REFERENCE Tascon,S., Burfelind,P.,
AUTHORS Burfelind,P.
TITLE Isolation and characterization of a novel human gene, NIF3L1, and
        its mouse ortholog, Nif3l1, highly conserved from bacteria to
        mammals
JOURNAL Cytogenet. Cell Genet. 90 (3-4), 330-336 (2000)
MEDLINE 20573864
PUBMED 11124544
REFERENCE 2 (bases 1 to 1425)
AUTHORS Tascon,S., Burfelind,P. and Engel,W.
TITLE Direct Submision
JOURNAL Submitted (29-JUN-2000) Institute for Human Genetics, University of
        Goettingen, Heinrich-Dueker Weg 12, Goettingen 37073, Germany
FEATURES
        source
        1..1425
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /chromosome="2"
        47..1180
        /codon_start=1
        /product="NIF3L1 protein"
        /protein_id="AA044846.1"
        /db_xref="GI:12006403"
        /translation="MSSGVRPVPTVRFPVDSLICNRSRPFMDIKALLSLNDFASL
        PASMNDNVLVEPSPRYNTVLTFLNDLTERVMEVLOKADILSYRPIFRPKR
        ITWNTKRLVLRALENRGITISPHATDAPGVNMLAKGLGACTSRPIHPSKAP
        MYPLEGHNRHVEFNNTYODLDKWSAVKSIDGVSVTSFSAKNEQTRINLNCOT
        MOVDLFSRNKOLQOKTILSELKPLILMTGRLCTLDESVSLAMIDRIKRLHSL
        RLALAGVRLKLEISQVKAACAGSGSVLQGVADLYLTGEMSHHDTLDAASOGIN
        ILCEHSNTERGFI.SDLRMLDSLEKINILITSETRDLPLVY"
BASE COUNT      385 a      331 c      314 g      395 t
ORIGIN
Alignment Scores:
Pred. No.:      6.41e-146      Length:      1425
Score:          1799.00      Matches:      350
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             9      Gaps:      0

```



Score: 1799.00 Matches: 350  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-745-506-37 (1-350) x BC007654 (1-1440)

QY 1 MetAplLeuAlaLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20  
 DB 118 ATGGATTGAAAGGCTCTCTTCTCTTCTGAAAGATCTTGCACTCCCTCTCTTGGTGAG 177  
 QY 21 SerTPAspAsnValGlyLeuLeuValGluProSerProProHisThrValAsnThrLeu 40  
 DB 178 AGTTGGACAAATGTTGATTACTGCTGGAAACCAAGCCCAACATCTGTAAATACACTC 237  
 QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluLeuGluValAsp 60  
 DB 238 TTCCCTGACCACATGACCTGACAGCAAGATGAGAGAGTGTCTCCAAAGAGGACAGAC 297  
 QY 61 LeuAlaLeuSerThrHisProProIlePheArgProMetLeuArgIleThrTrpAsnThr 80  
 DB 298 GTCATCTCTCTCTACCATCCGCTATCTTCCGACCATGAAGCGCTGAACCTGGAAACACA 357  
 QY 81 TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleThrSerProHis 100  
 DB 358 TGGAAAGAGCGCTGGTGATCCGGGCTCTGGAAACAGAGTGGTATCTACTCTCTCAT 417  
 QY 101 ThrAlaTyrAspAlaAlaProGluGlyValAsnAsnTrpLeuAlaLysGlyLeuGlyAla 120  
 DB 418 ACAGCCTATGATGCTGGCCGCCAGGCGCTCAACACTGGTGGCTAAAGGCTTGGAGCT 477  
 QY 121 CysThrSerArgProIleHisProSerLysAlaProAsnTrpProThrGluGluAsnHis 140  
 DB 478 TGTACCTCCAGCGCCATACATCTCTCAAGAGCTCCCACTACCTACAGAGGAAACACAC 537  
 QY 141 ArgValGluPheAsnValAsnTrpThrGluAspLeuAspLysValMetSerAlaValLys 160  
 DB 538 CGAGTGAATTCACAGCTTAACATACACCAAGACCTGGCAAAAGTCAATGCTCGAGTGA 597  
 QY 161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGluGlnThr 180  
 DB 598 GGAATGACGGGTCT 657  
 QY 181 ArgIleAsnLeuAsnCysThrGluLysAlaLeuMetGluValValAspPheLeuSerArg 200  
 DB 658 CGGATTATCTGATTTGATCTACAGAGGCTTTGATCTAGAGTGTGATTTCTTTCCCG 717  
 QY 201 AsnLysGluLeuTyrGluLysThrGluLeuSerLeuGluLysProLeuLeuHis 220  
 DB 718 AACCAACACTTATATCAGAAAGAGAAATCTGTCTACTGAGAGAGCTTGGCTTCTACAT 777  
 QY 221 ThrGlyMetGlyArgLeuGlyThrLeuAspGluSerValSerLeuAlaThrMetLeuAsp 240  
 DB 778 ACTGGATGAGGAGGCTTATGACACACATGATGATCTCTCTCTCTCTCTCTCTCTCT 837  
 QY 241 ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValGlyArgThr 260  
 DB 838 CCAATTAAGACACTTAATACTATCTCATTTCTGAGCTTTGGGGTGGGAGAACACC 897  
 QY 261 LeuGluSerGluValLysValAlaLeuCysAlaGlySerGlySerSerValLeuGln 280  
 DB 898 TTAGAGTCTCAAGTCAAAAGTGGGCCCTGTGTGCTGGTGGAGGAGCGCTTCGACG 957  
 QY 281 GlyValGluAlaAspLeuTyrLeuThrGlyGluMetSerHisHisAspThrLeuAspAla 300  
 DB 958 GGTGTGAGGCTGACCTTACCTACACAGGTGAGANTCTCCATCATATCTTTGGATGCT 1017  
 QY 301 AlaSerGluGlyLysAsnValIleLeuGlyGluHisSerAsnThrGluArgGlyPheLeu 320  
 DB 1018 GCTTCCCAAGATTAATATCTCATCTCTGTGAACACAGACACTGAAAGAGGCTTCTGT 1077  
 QY 321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleIleLeuSer 340

DB 1078 TCGACCTTCGAGATATGCTGATTTCTCACTTGTGAAATATATATATATATATCA 1137  
 QY 341 GluThrAspArgAspProLeuGluValVal 350  
 DB 1138 GAGACTGACAGGAGGACCTCTTCAGTGTGTA 1167

RESULT 5  
 BD158389 1574 bp DNA linear PAT 17-JAN-2003  
 LOCUS  
 DEFINITION  
 Primer for synthesizing full-length cDNA and use thereof.  
 ACCESSION  
 BD158389.1 GI:27864147  
 VERSION  
 JP 2002191363-A/13232.  
 KEYWORDS  
 Homo sapiens (human)  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 1574)  
 Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
 Primer for synthesizing full-length cDNA and use thereof  
 Patent: JP 2002191363-A 13232 09-JUL-2002;  
 HELIX RESEARCH INSTITUTE  
 OS Homo sapiens (human)  
 PN JP 2002191363-A/13232  
 PD 09-JUL-2002 JP 2002280990  
 PE 28-JUL-2000 JP 2002280990  
 PI TOSHIO OTA,TAKAO ISOGAI,TEISUO NISHIKAWA,KOJI HAYASHI,KAORU  
 PI SAITO,  
 PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,  
 PI KEIICHI NAGAI,TEISUJI OTSUKI  
 PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC  
 10,  
 PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC  
 Primer for synthesizing full-length cDNA and use thereof FH key

FEATURES  
 source FT CDS  
 1..1574 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 420 a 361 c 372 g 421 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 7.24e-146 Length: 1574  
 Score: 1799.00 Matches: 350  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-745-506-37 (1-350) x BD158389 (1-1574)

QY 1 MetAplLeuAlaLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20  
 DB 271 ATGGATTGAAAGGCTCTCTTCTCTTCTGAAAGATCTTGCACTCCCTCTCTTGGTGAG 330  
 QY 21 SerTPAspAsnValGlyLeuLeuValGluProSerProProHisThrValAsnThrLeu 40  
 DB 331 AGTTGGACAAATGTTGATTACTGCTGGAAACCAAGCCCAACATCTGTAAATACACTC 390  
 QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluLeuGluValAsp 60  
 DB 391 TTCCCTGACCACATGTTGATTACTGCTGGAAACCAAGCCCAACATCTGTAAATACACTC 450  
 QY 61 LeuAlaLeuSerThrHisProProIlePheArgProMetLeuArgIleThrTrpAsnThr 80  
 DB 451 CTCATCTCTCTTACCATCCGCTATCTTCCGACCATGAAGCGCTAACTGGAACACA 510

QY 81 Trrpysgluargleuvalilleargalaleuclunsnargvalglylletyrserrprohis 100  
 DB 511 TGGAGGAGCGCCGCGTGATCCGGGCTCTGGAGAACAGAGCGGTATCTACTCTCCAT 570  
 QY 101 Thrilatrraspalaalaproglnglyvalasnsntrpleuvalaleysglyleuclvala 120  
 DB 571 ACACGCTATGTGCTGCGCCCGCCAGGGCTCAACAACCTGCTTGCTTAAGGCTTGAGACT 630  
 QY 121 CysThrserarproilehisproserlysalaoproasnryrprothrglucllyasnhis 140  
 DB 631 TGTACCTCCAGGCCCATCATCTCTCCAAAGCTCCCAACTACCTACAGAGGAAACAC 690  
 QY 141 Argvalglupheasnvalasnryrthrclnaspleuaplyvalmetseralavallys 160  
 DB 691 CGAGTAGAATTCACAGCTTACTACACCCAGACCTGGCAAGATCATCTGCACTGAAA 750  
 QY 161 Glyileaspglyvalservalthrserpheserilaargthrnglyasngluglphr 180  
 DB 751 GGAATGTACGGTGTCTGTCACTCTTTCTGTCTAGAGCTGTAATGAGAGAACAAACA 810  
 QY 181 ArgileasnleuasnrcysThrclnlysalaleuclunsnargvalaspheleuserarg 200  
 DB 811 CGGATTAATCTGAATTTGTAACAGAGCTTTGATGAGAGGTAGATTTCTTCCGG 870  
 QY 201 AsnlysglnleuTyrglnlyThrGluileuserleuclulysproleuclunhis 220  
 DB 871 AACAAACAATCTTATCAGAGACGAAATCTGTCTACGAGAGACCTTCTCTTACAT 930  
 QY 221 ThrGlymetgllyargleucysThrleuaspGlnserValserleuAlaThrMetleasp 240  
 DB 931 ACTGGAATGGAGCGGTATTCACACACTGATGATGTCTCTCTGCAACATGATTAT 990  
 QY 241 ArgileuarghlsleuclunserhisileargleuAlaLeucllyvalglyrthr 260  
 DB 991 CGAATTAACAAACACACCTAAACTATCTCATATTCCTTACGCTTGAGCGGTGGAGAAC 1050  
 QY 261 Leucluserglnvallyvalalaleuclunsnargvalyserysergyservallleucln 280  
 DB 1051 TTAGAGCTCAAGTCAAGAGTGGCGCTGTGTCTGTCTGTGGAGAGCGCTTCTGAG 1110  
 QY 281 GlyvalglualaspLeuTyrlleuThrGlyglumetserhisaspThrleuaspala 300  
 DB 1111 GGTGTGAGGCTGACCTTACTCTCACAGTGAATGATCCCATCATGATCTTGATGCT 1170  
 QY 301 AlaserGlnlyleasnvalilleuclunserasnryrthrGluarglypheleu 320  
 DB 1171 GCTTCCCAAGAAATTAATGATCTCTGTGTGACACACACACAGAGAGGCTTCTT 1230  
 QY 321 SeraspLeuargaspmetleuasnserhisleuclunsnargvalyserysergyservallleucln 340  
 DB 1231 TCTGACCTTCGAGATATCTGATCTCTGATCTGAGATTAAGATTAATTAATCTCTATCA 1290  
 QY 341 GluThraspargaspProleuclunval 350  
 DB 1291 GAGACTGACAGGAGACCTCTTCAGGTGTA 1320  
 RESULT 6  
 AK023378  
 LOCUS Homo sapiens cDNA FLJ13316 f1s, clone OVARC1001555, weakly similar  
 DEFINITION to NGG1-INTERACTING FACTOR 3.  
 ACCESSION AK023378  
 VERSION AK023378.1 GI:10435289  
 KEYWORDS oligo cloning; f1s (full insert sequence).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE  
 AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,  
 Nishikawa,T., Nagai,K., Sugano,S., Shikatori,A., Sudo,H.,  
 Matsushima,M., Hosokiri,T., Kaku,Y., Kodaira,H., Kondo,H.,  
 Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,

ONO, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,  
 Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A.,  
 Nakamura, Y., Nagahata, K., Masuko, Y., Niimiya, K. and Iwagami, T.  
 MEDO human cDNA sequencing project  
 Unpublished  
 2 (bases 1 to 1574)  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS Isogai, T. and Otsuki, T.  
 TITLE  
 JOURNAL  
 COMMENT  
 Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,  
 Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
 MEDO human cDNA sequencing project supported by Ministry of  
 International Trade and Industry of Japan, cDNA full insert  
 sequencing. Research Association for Biotechnology: cDNA library  
 construction, 5'- & 3'-end one pass sequencing and clone selection:  
 Helix Research Institute (supported by Japan Key Technology Center  
 etc.) and Department of Virology, Institute of Medical Science,  
 University of Tokyo.  
 FEATURES  
 source  
 1..1574  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="OVARC1001555"  
 /tissue\_type="ovary, tumor tissue"  
 /clone\_lib="OVARC1"  
 /note="Cloning vector: pME18SFL3"  
 271..1323  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="BAB14551.1"  
 /db\_xref="GI:10435290"  
 /translation="MDKLALISLNDPRLSFLSDWNVGLVPPSPHYVNTLPLTN  
 DLTEVEVEVLOKRAADLISYHPPIFRMKRTIWNKRLVIALENRQVYSPTA  
 YDAPOGVNMLLAGLACTSRPIHPSKAPYPTGHNRYVFNQDLDKVSAYK  
 GIDGVSFSGARTGNEQRTIRINOCALMOYVDFSRKOLYOTELISLKPPL  
 LHTGMRCTCTDEVSILATMDRIKRHKIISHILAIAGVGRLESOYKVALGAGSS  
 SYLVGVEADLITGEMSHHDTLDAASGIVNITLDEHSNTENGFLSLRDMDSLENK  
 INILSERDRLPQV"  
 BASE COUNT 420 a 361 c 372 g 421 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7,24e-146 Length: 1574  
 Score: 1799.00 Matches: 350  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 9  
 US-09-745-506-37 (1-350) x AK023378 (1-1574)  
 QY 1 MetaspLeuysAlaleuclunserleuasnasphealaSerleuSerPheAlaGlu 20  
 DB 271 ATGATTTGAAGGCTCTCTCTTCTCTTGAATGATCTTGATCCCTCTGCTTGCTGAG 330  
 QY 21 SerTrpaspasnvalglyleuclunvalGluProserProProhisThvalasnThreu 40  
 DB 331 AGTGGGCAATGTTGGATTCTGTGACACCAACCCACACACATACATGTAATACACTG 390  
 QY 41 PheleuThrAsnaspLeuThrGlnGluValmetGluValleuGlnlysalasp 60  
 DB 391 TTCTGACCAATGACCTACTGAGAGATGATGGAGGAGGTGCTGCAAAAGAGCAGAC 450  
 QY 61 LeuileuSerTyrlHisProProlePheargProMetlysaGlyleThrTrpAsnThr 80  
 DB 451 CTCATTTCTCTTACACCGGCTATCTTCGAGACCAAGAGGATTAACCTGGAACACA 510  
 QY 81 Trrpysgluargleuvalilleargalaleuclunsnargvalglylletyrserrprohis 100  
 DB 511 TGGAGGAGCGCGCTGATCCGGGCTCTGGAGAACAGAGCGGTATCTACTCTCCAT 570  
 QY 101 Thrilatrraspalaalaproglnglyvalasnsntrpleuvalaleysglyleuclvala 120

```

Db      571 ACAGCCTATGATGTCGCCGCCAGGCGTCAACAACATGCTGGCTAAAGGCTTGAGCT 630
Qy      121 CysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGlyAsnHis 140
Db      631 TGTACCTCCAGGCCCATCATCTCTTCCAAAGCTCCCACTACCCTTCACAGAGGAAACAC 690
Qy      141 ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaValys 160
Db      691 CCAGTAGAATTCACGCTTACTACACCCAGACCTGGACAAAGTCTGTCGACAGTGAAG 750
Qy      161 GilyleaspglyvalserValThrSerPheSerAlaArgThrGlyAsnGluGlnThr 180
Db      751 GGAATTTAGACGGTGTCTCTACTCTTTCTTCTGTTAGACATCGTAAATGAGAAACA 810
Qy      181 ArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGlnValAlaAspHisLeuSerArg 200
Db      811 CGGATTTAATCTGAATTTGTCTCAGAAAGCGCTTGTATGACGTTAGATTTTCCTTCCGG 870
Qy      201 AsnLysGlnLeuTyrGlnLysThrGluIleLeuSerLeuGluLysProLeuLeuHis 220
Db      871 AACAAACAATTTATTCAGAAAGCGAAATTCCTCTCCTGACAGACCTTTCCTTCTACAT 930
Qy      221 ThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetIleasp 240
Db      931 ACTGGATGGAGCGGTTATGACACACTGATGATCTGCTCCCTGGCAACCATGATTGAT 990
Qy      241 ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyAlaGlyThr 260
Db      991 CGAATTTAAAGACACCTTAAACATCTCATTTGCTTACGCTTGGGGGGGGAAGAAC 1050
Qy      261 LeuGluSerGlnValLysValAlaAlaLeuCysAlaGlySerGlySerSerValLeuGln 280
Db      1051 TTAGAGTCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1110
Qy      281 GlyValGluAlaAspLeuTyrLeuThrGlyGluMetSerHisHisAspThrLeuAspAla 300
Db      1111 GGTTGTGAGCGTACCTTACCTCAGAGTGAATCTCCATCATCATCTTTGGATGCT 1170
Qy      301 AlaSerGlnGlyIleAsnValIleLeuGlyHisSerAspThrLysGlyGlyPheLeu 320
Db      1171 GCTTCCCAAGAAATTAATGTCATCTCTGGAACAGCAACCTAACAGAGCTTTCCT 1230
Qy      321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleIleLeuSer 340
Db      1231 TCTGACCTTCGAGATATGCTGATTTCTCACTTGGAAATTAATATATATATCTATCA 1290
Qy      341 GluThrAspArgAspProLeuGlnVal 350
Db      1291 GAGACTGACAGGAGACCTCTTCAGTGGTA 1320

RESULT 7
AB038949 1606 bp mRNA linear PRI 23-FEB-2001
LOCUS      Homo sapiens ALS2CR1 mRNA for amyotrophic lateral sclerosis 2,
DEFINITION candidate 1, complete cds.
ACCESSION  AB038949
VERSION     AB038949.1 GI:12862477
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)
REFERENCE   Hadano,S., Yanagisawa,Y., Skaug,J., Fichter,K., Nasir,J.,
AUTHORS     Marindale,D., Koop,B.F., Scherer,S.W., Nicholson,D.W.,
            Rouleau,G.A., Ikeda,J.-E. and Hayden,M.R.
TITLE       Cloning and characterization of three novel genes, ALS2CR1,
            ALS2CR2, and ALS2CR3, in the juvenile amyotrophic lateral sclerosis
            (ALS2) critical region at chromosome 2q33-q34: candidate genes for
            ALS2
JOURNAL     Genomics 71 (2), 200-213 (2001)
MEDLINE     21100893

```

```

PUBMED 11161814
REFERENCE 2 (bases 1 to 1606)
AUTHORS   Hadano,S., Ikeda,J. and Hayden,M.R.
TITLE     Direct Submission
JOURNAL   Submitted (24-FEB-2000) Shinji Hadano, Tokai University, The
            Institute of Medical Sciences; Bohseldai, Isehara, Kanagawa
            259-1193, Japan (E-mail:shinji@eng.med.u-tokai.ac.jp,
            Tel:+81-463-91-5095, Fax:+81-463-91-4993)
FEATURES
Source
1..1606
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q33"
/note="IMAGE:32584-UniGene ID: Rs.21943"
1..1606
/gene="ALS2CR1"
286..1338
/gene="ALS2CR1"
/note="candidate 1"
/codon_start=1
/product="amyotrophic lateral sclerosis 2"
/protein_id="BAB32499.1"
/db_xref="GI:12862478"
/translation="MDKALSLSDNPFASLFAESMDNGLVPEPPHTVNTLFLTN
DLEFVMEVYLOKKADLLISYHPFRPMKRTWMTKEVTVRALENRGVSPHDA
YDAAPGVNWMFLAKGLACTSRPIHRSKAPNTYEGNHRREPNTODDLYMSAVK
GIDGVSTFSARTGNEBQRIKNTOKAMQVDFPLSKKOLYKRTLEKPL
LHTGMGRLTLDSEVSLATMIDRIKHLKLSHRLALGVGRILESOVKVVALCAGGS
SVLGVEADLYLTGEMSHDPTDPAASQGINVILCHSNTNERGPLSDLRDLJSHLENK
INILSETRDPDQVY"
1566..1571
/gene="ALS2CR1"
BASE COUNT 438 a 365 c 375 g 428 t
ORIGIN
Alignment Scores:
Pred. No.: 7,42e-146 Length: 1606
Score: 1799.00 Matches: 350
Percent Similarity: 100.008 Conservative: 0
Best Local Similarity: 100.008 Mismatches: 0
Query Match: 9 100.00% Indels: 0 Gaps: 0
US-09-745-506-37 (1-350) x AB038949 (1-1606)
Qy      1 MetAspLeuLysAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
Db      286 ATGGAATTTGAAGCGTCTCTTCTGCTGAATGACTTTGCATCCCTCGTTGGTGGAG 345
Qy      21 SerThrAspAsnValGlyLeuLeuValGluProSerProThrHisThrValAsnThrLeu 40
Db      346 AGTTGGACAAATTTGGATTATCGTGGGAACAAGCCACACATATCTGTAATATACATC 405
Qy      41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGlnLysAlaAsp 60
Db      406 TTCCTGACCAATACCTGACTGAGGAAGTATGAGAGAGTGGCTGCAAAAGAGCGAGAC 465
Qy      61 LeuIleLeuSerTyrHisProIlePheArgProMetLysArgIleThrPasnThr 80
Db      466 CTGATTCCTCTCCATACATCGCGCTATCTTCGACCAACATGAAGCGCATTAACCTGGAACA 525
Qy      81 ThrPheGluArgLeuValIleArgAlaLeuGluAsnArgValAlaGlyIleTyrSerProHis 100
Db      526 TGGAAAGAGCGCTGTGTGATCGCGCTGTGGAAGAGACTCGGTATCTACTCTCCAT 565
Qy      101 ThrAlaTyrAspAlaAlaProGlnGlyValAsnAsnThrLeuAlaLysGlyLeuGlyAla 120
Db      586 ACAGCCTATGATGTCGCCGCCAGGCGTCAACAACATGTTGGCTAAAGGCTTGAGCT 645
Qy      121 CysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGlyAsnHis 140

```

Db	646	TGATCCTCCAGCCCATACATCTTCCAAAGCTCCCACTACACCTACAGAGGAAACCC	705	/cell_type="hematopoietic stem/progenitor cells"
Qy	141	ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaValLys	160	/note="Isolated from a Myelodysplastic Syndrome patient"
Db	706	CGAGTAGATTCACAGCTTAACCTACCCAAAGACCTGGACAAAGTATGTCTGCAGTAA	765	1. 1579
Qy	161	GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGlnThr	180	/gene="MDS015"
Db	766	GGAATGACGGGTGTTTCTGTCACTCTTTCTGCTAGACCTGGTAATAGAGAACAAAC	825	/gene="MDS015"
Qy	181	ArgIleAsnLeuAsnGlyThrGlnLysAlaLeuMetGlnValAlaAspPheLeuSerArg	200	/gene="MDS015"
Db	826	CGGATTAATTCGAAATGTACTCAGAAAGGCTTTGATGAGGCTGATGATTTCTTTCCGG	885	/codon_start=1
Qy	201	AsnLysGlnLeuTyrGlnLysThrGluIleLeuSerLeuGluLysProLeuLeuLeuHis	220	/product="MDS015"
Db	886	AACAACACACTTTATACAGAACACGAAATTTCTGTACAGGAAACCTTGTCTTACAT	945	/protein_id="AAG14952.1"
Qy	221	ThrIleMetGlyArgLeuGlyThrLeuAspGluSerValSerLeuAlaThrMetIleAsp	240	/db_xref="GI:10197632"
Db	946	ACTGGAATGGACGGTATATGACACAGTGAATCTGTCTCCCTGCAACCATGATGAT	1005	/translation="MDKLKLLSLINDPASTSPASEMDNVGLVPEPPHTVNTFLTN
Qy	241	ArgIleLysArgHisIleLeuLysLeuSerHisIleArgGluAlaLeuGlyValGlyArgThr	260	DTEVEEVLQKKADLLSTHPPIFRPMKRLNLEHMKRELVLRALENRGITSPHT
Db	1006	CGAATAAAAGACACTAAACTATCTCATATTCCTTAGCCCTTGGGGGTGGGAGAAC	1065	AYDAPQGVNVLAKGLGACTSRPIHPSKAPNPTGSHRELVNNTQDDLVKASAV
Qy	261	LeuGluSerGlnValLysValAlaLeuGlyAlaGlySerGlySerSerValLeuGln	280	KGIDGVSTFSFASRQNEQRLNINCQKALMQVDFLSRNKQLYQRTLLSEKPL
Db	1066	TTAGAGTCTCAAGTCAAGAGCTGTGGCCCTGTGTGTGTCTGTGGAGACGCTTGTGAG	1125	LHTGEMGLCTLDESVSLATMIDIRIKRLKLSHRLALGVRTLESQYKVAACAGS
Qy	281	GlyValGluAlaAspLeuTyrLeuThrGlyGluMetSerHisAspThrLeuAspAla	300	SSVGVGVADLYLIGEMSHDPTLDAASGIVNVILCEHSNTERGFLSLDRMDSHLEN
Db	1126	GGTGTGGAGCGTACCTTTACCTCACAGTGGATGATGCCATCATGATATCTTGGATGCT	1185	KINILSTDDPQVYV"
Qy	301	AlaSerGlnGlyIleAsnValIleLeuGlySerGlnHisSerAsnThrGluArgGlyPheLeu	320	
Db	1186	GCTTCCCAAGAAATTAATGTCATCTCTGTGTGACACAGACAACTGAAGAGGCTTCT	1245	
Qy	321	SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleLeuSer	340	
Db	1246	TCTGACCTTCGAGATATCTGGATCTCTCAGTGGAGATTAAGATTAATATCTTCTATCA	1305	
Qy	341	GluThrAspArgAspProLeuGlnValVal	350	
Db	1306	GAGACTGACAGGACCTCTTCAGGTGTGA	1335	
RESULT 8				
AF182416		1579 bp	mRNA	linear
LOCUS	AF182416			
DEFINITION	Homo sapiens MDS015 (MDS015) mRNA, complete cds.			
ACCESSION	AF182416.1	GI:10197631		
VERSION				
KEYWORDS				
SOURCE				
ORGANISM	Homo sapiens (human)			
REFERENCE				
AUTHORS	Huang C., Qian B., Tu Y., Gu W., Wang Y., Han Z. and Chen Z.			
TITLE	1 (bases 1 to 1579)			
JOURNAL	Unpublished			
FEATURES				
source				
1. 1579				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				



OY		220	HistHrGlyMetGLyATgTgCysThLeuaspLusSerValSertLeuAlaEThrMeTle	239
Dd		904	CATACTGGAAATGGACCGTTATGCACACTGGATGAATCTGTCTCCAGGCCAACCATATT	963
OY		240	AspaRgileLySaRgHisLeuLySLeuSerHisIleaRgLeuAlaLeuGlyValGIyARg	259
Dd		964	GATCGAATATAAAAAACACACCTAAACCTAATTCATATTCGCTTAGCCCTTGAGGGGGAGA	1023
OY		260	ThreugLusSerGlnVallySValValAlaLeucysAlaGlySerGlySenSerValLeu	279
Dd		1024	ACCTTAGAGTCTCAAGTCATAAAGTGTGGCCCTGTGTCTGTGTGTGGAGACGGCTTG	1083
OY		280	GInGlyValGIuaLaasPLeuTyLeuThrcylGluMetSerHisHSAsPyThLeuASP	299
Dd		1084	CAGSGTGTGGAGCGTACCCTTACCTCACAGGTGAGATGTCATCATGAACTTGGAT	1143
OY		300	AlaAlaSerGlnGlyIleasnValIleLeucysGlnHSerAsnThrcyluaRglyPhe	319
Dd		1144	GCTCCTTCGCCAAGAAATAATGTCATCTCTGTGAACACACCAACTGAACGAGGCTTT	1203
OY		320	LeuSerAsPLeuAUGAsPMeLeuASerHisISLeuGluasnLySIleAsnIleIleu	339
Dd		1204	CTTCTGACCTTCACAGATATCTGTGATTTCTACCTGGAGAAATAAGATAATATTATCTTA	1263
OY		340	SerGluThraSParGaSPProLeuGlnValVal	350
Dd		1264	TCAGAGACTGACAGAGGCCCTCTCAGAGTGSTA	1296
RESULT 9				
LSHM805522				
LOCUS	HSN805522	1353 bp	mRNA	linear PRI 12-JUL-2002
DEFINITION	Homo sapiens mRNA; cDNA DKFPZp62L015 (from clone DKFPz62L015).			
ACCESSION	AL834430			
VERSION	AL834430.1 GI:21740154			
KEYWORDS				
ORGANISM	Homo sapiens (human)			
AUTHORS	Mumaiyola, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. Bloeker, H., Boecker, M., Brandt, P., Mewes, H.W., Well, B. and Wiemann, S.			
TITLE	Direct Submission			
JOURNAL	Submitted (09-JUL-2002) 1. D-85764 Neuberberg, GERMANY			
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Sequenced by GBF (National Research Centre for Biotechnology Ltd., Bundeschweigen/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFPZp62L015) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/. Location/Qualifiers			
FEATURES				
source	1..1353			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="DKFPZp62L015"			
	/tissue_type="melanoma (Memo cell line)"			
	/clone_lib="762 (synonym: hme12). Vector pSPORT1; host DH10B; sites NotI + SalI"			
	/dex_stage="adult"			
	128..1233			
	polya_signal			
	polya_site			
BASE COUNT	457 a 283 c 280 g 333 t			
ORIGIN	1252			
Alignment Scores:				
Pred. No.:	1.07e-138	Length:	1353	
Score:	1715.00	Matches:	332	

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	95.33%	Indels:	0
DB:	9	Gaps:	0

US-09-745-506-37 (1-350) x HSM805522 (1-1353)

QY	19	AlaGluSerTrpAspAsnValGlyLeuLeuValGluProSerProHisIleThrValAsn	38
DB	2	GCTGAGAGTTGGACATATGTTGGATTACTGGTGGACCAAGCCACCATACTGTAAT	61
QY	39	ThrLeuPheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGlnLys	58
DB	62	ACACTCTTCCTGCACCAATGACCTGACTGAGAAAGTGAAGAGAGAGTCTCAAAAGAG	121
QY	59	AlaAspLeuIleLeuSerTrpHisProProIlePheIleArgProMetLysArgIleThrTrp	78
DB	122	GCAACCTCTATCTCTCTCAACCTCCGCTATCTTCCGACCCATGACGATTAACCTGG	181
QY	79	AsnThrTrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSer	98
DB	182	AACACATGGAAGAGGCGCTGGTGATCGGGCTCGAGAAACAGATGCGATCTACTCT	244
QY	99	ProHisThrLarTrpAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeu	118
DB	242	CCTCATCACACCTATGATGTGTGGCCCCGAGGCGTCAACATCTGTTGGCTAAAGGGCTT	301
QY	119	GlyAlaCysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGly	138
DB	302	GGACCTTGATCTCCAGGCCCATACTCTTCCAAAGCTCCCACTTACCCTTACAGAGGA	361
QY	139	AsnHisArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAla	158
DB	362	AACACCGAGTATCAATTCACCGTTACTACACCCAAAGACCTGACCAAGTCAATGCTGCA	421
QY	159	ValLysGlyIleAspGlyValSerValIleIleSerPheSerAlaArgThrGlyAsnGluGlu	178
DB	422	GTCGAAGCAATTCACGGTGTCTTCTGCACCTCTTTTCTGCTAGGACGGTAATGAGAA	481
QY	179	GlnThrArgIleAsnLeuAsnCysTrpGlnLysAlaLeuMetGlnValValAspPheLeu	198
DB	482	CAAACACGAGTATATCTGAAATTTACTACGAAAGCGTTTGATGACAGTGTGATTTCTT	541
QY	199	SerArgAsnLysGlnLeuTyrGlnLysThrGluIleLeuSerLeuGlnLysProLeuLeu	218
DB	542	TCCCGGAACAAACAACCTTATATCAGAAAGCGAAATTCGTCACTGAAACACCTTGGCTT	601
QY	219	LeuHisThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMet	238
DB	602	CTACATACTGGAATGGGACGGTATATGACACTGGATGAATCTGTCTCCCTGGCAACATG	661
QY	239	IleAspArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValGly	258
DB	662	ATTGATGCAATTAATAAAGACACCTTAACATATCTCATATTCGCTTGACCTTGGGGTGGG	721
QY	259	ArgThrLeuLeuSerGlnValLysValAlaLeuCysAlaGlySerGlySerSerVal	278
DB	722	AGAACCTTAAAGTCTCAAGTCAAAAGTCGTGGCCCTGTGGTGGTTCGGGACAGCGCTT	781
QY	279	LeuGlnGlyAlaGluAlaAspLeuTyrLeuThrGlnGlyLysMetSerHisAspThrLeu	298
DB	782	CTGAGGGGTGTAGGCTGACCTTTACTCTACACAGGAGATGTCCATCAATGATCTTGG	841
QY	299	AspAlaAlaSerGlnGlyIleAsnValIleLeuCysGluHisSerAsnThrGluArgGly	318
DB	842	GATGCTGCTTCCCAAGGAATTAATGTCATCTCTGTGAACACAGCAACACTGAACGAGGC	901
QY	319	PheLeuSerAspLeuLysArgMetLeuAspSerHisIleGluGluAsnTyrIleAsnIleIle	338
DB	902	TTTCTTTCTGACTTGGAGATATGCGGTCTTCACTTGGAGAAATAGATTAATTTATCT	961
QY	339	LeuSerGluThrAspArgAspProLeuGlnValVal	350

Db 962 CTATCAGAGACTGACAGGACCTCTTCAGGTGTA 997

RESULT 10  
AF284439 1836 bp mRNA linear ROD 16-JAN-2001

LOCUS AF284439  
DEFINITION Mus musculus NIF3L1 mRNA, complete cds.  
ACCESSION AF284439  
VERSION AF284439.1 GI:12034693

KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS Tascou, S., Uedelhoven, J., Dikens, C., Nayeria, K., Engel, W. and  
Burfeind, P.  
TITLE Isolation and characterization of a novel human gene, NIF3L1, and  
its mouse ortholog, Nif3l1, highly conserved from bacteria to  
mammals

JOURNAL  
MEDLINE Cytogenet. Cell Genet. 90 (3-4), 330-336 (2000)  
PUBMED 1124544

REFERENCE  
2 (bases 1 to 1836)  
Tascou, S., Burfeind, P. and Engel, W.  
AUTHORS Direct Submission  
JOURNAL Submitted (02-JUL-2000) University of Göttingen, Institute of Human  
Genetics, Heinrich-Duker weg 12, Göttingen 37073, Germany

FEATURES  
source  
1..1836  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/chromosome="1"  
508..1638  
/codon\_start=1  
/product="NIF3L1"  
/protein\_id="AA045961.1"  
/db\_xref="GI:12034694"  
/translation="MLSRAHIVPTSVQRAQSWICRSRSEFNDKALISSLNDPFLSST  
AESWDNGLVPEPSPHTVNTLFLTNDTEVMDALQKADFLLSHPIFRPKHT  
TWKTKBCLVIRALENRYAVSPHTAYDAPOGVSNWAKGLGCTTRPIHPSRAVDY  
PTGEAHLRFVSNSSODLDKRVSTLRGGSVTSFPRCGEQRTRSLNCTOKLM  
OVLAFISODRLOYKTEITSLKPLILHTGRLCTDESVLSIMIRITHKLSH  
LRALIGVGTESQVKKYVALCAGSGSVLQGVADYLTGMSHHVDVDAASKGINVI  
LCEHSNTERGSLSEQLQEMLVGHPKKNITILSETRDPLRVV"

BASE COUNT 462 a 517 c 450 g 407 t

ORIGIN

Alignment Scores:  
Pred. No.: 1.89e-126 Length: 1836  
Score: 1575.00 Matches: 302  
Percent Similarity: 93.14% Conservative: 24  
Best Local Similarity: 86.29% Mismatches: 24  
Query Match: 87.55% Indels: 0  
Gaps: 0

US-09-745-506-37 (1-350) x AF284439 (1-1836)

QY 1 MetaspLeuYsAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20  
Db 586 ATGATCTGGAAGGCTCTCTCGCTTGATGATGATGATGATGATGATGATGATGATG 645

QY 21 SerTrpAspAsnValGlyLeuLeuValGluProSerProProHisThrValAsnThrLeu 40  
Db 646 AGCTGGGACAAATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 705

QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluHisLysAlaAsp 60  
Db 706 TTCTGACCAATGACCTGACGAGAGAGATGACGAGGCTCTGCCAAGAGAGACAC 765

QY 61 LeuLeuSerLeuTrpHisProProLeuPheArgProMetLysArgIleThrTrpAsnThr 80  
Db 766 TTCATTCCTCTACACATCACCTATTTCCGGCCCATGAGACACATTAATTGGAAACC 825

QY 81 TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTrpSerProHis 100  
Db 826 TGAAGAGGTCTCTCGTATCGCGGCTCTGGAGAAACAGAGTGCCTTACTCTCCACC 885

QY 101 ThrAlaTrpAspAlaIleAlaProGlnGlyAlaAsnSerProLeuAlaLysGlyAla 120  
Db 886 ACACCTATGATGAGGCGACCCAGGAGGTCAACAGCTGGTTGGCCAAAGGGCTTGGAACT 945

QY 121 CysThrSerArgProIleHisProSerLysAlaProAsnTrpProThrGluGlyAsnHis 140  
Db 946 TGCACTACACAGCCCATCCACCTTCCAGACTCCACAGCTCCACACAGAGAGGCTCAC 1005

QY 141 ArgValGluPheAsnValAsnTrpThrGlnAspLeuAspLysValMetSerAlaValLys 160  
Db 1006 CGACTAGAAATTCAGTGTGACACCGACGACCAACCTGGCAAAAGTCATGTCACTGAGA 1065

QY 161 GlyIleAspGlyLysSerValThrSerPheSerIleArgThrGlyAsnGluGluThr 180  
Db 1066 GGGGTGAGAGGTCTCTCTGTCATCTTTCTTCTGCGAGGTGATGATGATGATGATGATG 1125

QY 181 ArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGluValAlaAspPheLeuSerArg 200  
Db 1126 CGGATCAGCCTGAATGTACTACAGAACTTTGATGACGAGTGTAGCTTCTTCTCCAG 1185

QY 201 AsnLysGlnLeuTrpGlnLysThrGluIleLeuSerLeuGluLysProLeuLeuHis 220  
Db 1186 GACAGACAACTTTATCAGAAATGAAATTTCTTATGAGAAAGCCCTTCTCTGTCAT 1245

QY 221 ThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetIleAsp 240  
Db 1246 ACTGGAATGGAGACGGTTGTGCACACTGATGATGATGATGATGATGATGATGATGATG 1305

QY 241 ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValGlyArgThr 260  
Db 1306 CGAATCAAAACACACTTAACCTGCGCATCTCTGCTTACGCTTGGAGTGGGGAACAA 1365

QY 261 LeuGluSerGlnValLysValAlaLeuCysAlaGlySerGlySerValLeuGln 280  
Db 1366 TTAGAGTCCCAAGTCAAAAGTGTGGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1425

QY 281 GlyValGluAlaAspLeuTrpLeuThrGlyGluMetSerHisLysAspThrLeuAspAla 300  
Db 1426 GGAGTGGAGGCGGACCTTACTCAGAGTGAAGTCCACACAGATGATGATGATGATGATG 1485

QY 301 AlaSerGlnGlyIleAsnValIleLeuCysGluHisSerAsnThrGluArgGlyLeu 320  
Db 1486 GCTTCCAAAGGATCAATGTATCTTGTGTGAACACACACACACACACACACACACACAC 1545

QY 321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleLeuSer 340  
Db 1546 TCTGAGCTTCAGAAATGCTGGGTGTGATCTTTGGAATTAAGATTATATCTGTCT 1605

QY 341 GluThrAspArgAspProLeuGlnValAla 350  
Db 1606 GAGACAGACAGGACCTCTCGTGTGCTT 1635

RESULT 11  
AL645474 55374 bp DNA linear PRI 30-JAN-2002

LOCUS AL645474/C  
DEFINITION Human DNA sequence from clone RP5-92F19 on chromosome 1, complete  
sequence.

ACCESSION AL645474  
VERSION AL645474.10 GI:18476853  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Van Hellmond, Z.  
TITLE Direct Submission  
JOURNAL Submitted (30-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,

## COMMENT

Cambridge, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 On Feb 1, 2002 this sequence version replaced q1:18135182.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)  
 This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/BCP/Chr1>  
 RP5-925F19 is from the library RPc1-5 constructed by the group of Pierre de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pCIPAC2  
 IMPORTANT: This sequence is not the entire insert of clone RP5-925F19. It may be shorter because we sequence overlapping sections only once, except for a short overlap.  
 The true left end of clone RP11-478P15 is at 53375 in this sequence. The true right end of clone RP11-393M18 is at 2000 in this sequence.

## FEATURES

Location/Qualifiers  
 1..55374  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /clone="RP5-925F19"  
 /clone\_1lb="RPc1-5"  
 BASE COUNT 17426 a 9900 c 9927 g 18121 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 5.58e-73 Length: 55374  
 Score: 976.50 Matches: 232  
 Percent Similarity: 75.21% Conservative: 38  
 Best Local Similarity: 64.62% Mismatches: 71  
 Query Match: 54.28% Indels: 21  
 DB: 9 Gaps: 2

US-09-745-506-37 (1-350) x AL645474 (1-55374)

QY 1 MetAplleuYAlaleuSerleuAAspPheAlaSerleuSerPheAlaGlu 20  
 DB 8328 ATGATTTCAGAGGCTCTTCTGCTCTTGAATGACTTTCATCTCTGAG 8269  
 QY 21 SerTPASPAsnValGlyLeuValGluProSerProHisThrValAsnThrLeu 40  
 DB 8268 ACTTGTCAGCAATATGATTACTGTGGTGAACCAAGCATCATCTTAAACACAC 8209  
 QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGlnLysAlaAsp 60  
 DB 8208 TTCACAGACCAATGACTGAGTGAAGAGATGAGAGAGTCTGTGAAGAGACGAC 8149  
 QY 61 LeuLeuSerThrIleProPheArgProMetLysArgTleThrPaspThr 80  
 DB 8148 CTCATTCTCTCTGTCATCCACCAATTTTCAGACCTATTAAGGACATTAACCTA 8089  
 QY 81 TPlySGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleThrSerProHis 100

DB 8088 TGAAGAACTCTGCTGAATCCAGGCTGTCGAGAACGAACTGTATGATTTCTCCAC 8029  
 QY 101 ThrAlaTyrAspAlaAlaProGlnGlyValAsnAspThrLeuAlaLysGlyLeuAla 120  
 DB 8028 ATAGCTATGATCTCTCCACCCAGGAGTTTGCACGTGTGGCTTAAGGGCTGAGCT 7969  
 QY 121 CysThrSerArgProIleHisProSerLysAlaProAsnThrProThrGluGlyAsnHis 140  
 DB 7968 TGGACCTCTTAGCTGATGATCTTTCGAAAGCTTCCAAATGATCCATAGG----- 7918  
 QY 141 ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaVal-Ly 160  
 DB 7917 -----AACCC-ACCGAAGACCTGGACAAAGTACGCTGCGAGTGA 7877  
 QY 160 SGIYleAspGlyValSerValThrSerPheSerAlaValGlyGlyAsnGluGlnThr 180  
 DB 7876 AAGAGTTGCAATGTTCTGCTCTTGTCTGCGGATTTGATGATGATGAAAAAGAAC 7817  
 QY 180 ArgGluLeuAsnLeuAsnCysThrGlnLysAlaLeuMetGlnValAlaAspPheLeuSer 200  
 DB 7816 ATGACCTAGCTCTGATTAATCTCAGAACCTTTGATGACAGTGTGGCTTTCTCCCA 7757  
 QY 200 GAsnLysGlnLeuTyrGlnLysThrGlnLysLeuSerLeuGluLysProLeuLeuHis 220  
 DB 7756 GACCATTAATATGATCAGACAGACGTAATATATGTA-CTGGAGAACCTTGTCTTACA 7698  
 QY 220 sThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetLeu 240  
 DB 7697 TACTGGAATGGAGTGGTATGATCTACTGATGATTAATCTCCCTACAGCCATGACTGA 7638  
 QY 240 PArgIleLysArgHisLeuLysLeuSerHisIleatGlyLeuAlaLeuGlyValGlyArg 260  
 DB 7637 GTAAATCACAAGCCACTTAAACTATTAATATGATGCTTATGTTAGATACCGAGAC 7578  
 QY 260 rLeuGlu-SerGlnValLysValValAlaLeuCysAlaGlySerGlySerValLeuG 280  
 DB 7577 TGAAGAGTTCTAAAGCAAGGCTGGCCCTGTGCTGTGCTGGAGAGCGTCTGCG 7518  
 QY 280 InGlyValGluAlaAspLeuTyrLeu-Thr-----GlyGluMetSer 293  
 DB 7517 AGAGTTAGAGCTGACCTTTTACCTCCAGGGTGAAGGTCATCCAGTGGGATGCT 7458  
 QY 294 HisHisAspThrLeuAspAlaAlaSerGlnGlyIleAsnValIleLeuCysGlnHisSer 313  
 DB 7457 CATCATTAATATTTGATATGCTGCTCTTAAGAGCTAAATGCTATCTTTATGAATCAGC 7398  
 QY 314 AsnThrGluArgGlyPheLeuSerAspLeuArgAspMetLeuAspSerHisLeuGluAsn 333  
 DB 7397 AACACGTGAATTAACACTTCTTCTGACATTTGAGATATACTGGTCTTACTTGAAGAGT 7338  
 QY 334 LysIleAsnThrIleLeuSerGluThrAspArgAspProLeuGlnValVal 350  
 DB 7337 AAGATTAAATATATCTCT-TCAGAGACAAAGCAAGCACTTCTTATGTAGTA 7288

RESULT 12  
 BD149184 796 bp DNA linear PAT 17-JAN-2003  
 LOCUS DEFINITION  
 BD149184 Primer for synthesizing full-length cDNA and use thereof.  
 ACCESSION  
 BD149184.1 GI:27854942  
 VERSION  
 JP 2002191363-A/4027.  
 KEYWORDS  
 Homo sapiens (human)  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE  
 1 (bases 1 to 796)  
 Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
 Primer for synthesizing full-length cDNA and use thereof  
 JOURNAL  
 Patent: JP 2002191363-A 4027 09-JUL-2002;  
 HELIX RESEARCH INSTITUTE  
 OS Homo sapiens (human)

```

PN JP 2002191363-A/4027
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI MAGAI,TETSUITSU OTSUKI
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10,
PC C12P21/02,C1201/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1..796
/organism='Homo sapiens (human)'.
location/Qualifiers
1..796
/mol_type='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 192 a 204 c 202 g 195 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 2.76e-62 Length: 796
Score: 826.50 Matches: 170
Percent Similarity: 96.59% Conservative: 0
Best Local Similarity: 96.59% Mismatches: 6
Query Match: 45.94% Indels: 3
DB: 6 Gaps: 0
US-09-745-506-37 (1-350) x BD149184 (1-796)
QY 1 MetaspLeuYsAlaLeuLeuSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
DB 271 ATGGAATTGAAGGCTCTCTCTCTCTGTAATGATGCTTGATCCCTCTGCTGCTGAG 330
QY 21 SerTPAspAsnValGlyLeuLeuValGluProSerProProHisThValAsnThrLeu 40
DB 331 AGTGGGACATGTTGGATTACTGCTGGAGAACCAACCCACACATCTGTAATACACTC 390
QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluLysLysAlaAsp 60
DB 391 TTCTGACCATGACCTACTGAGAGAGATGATGGAGGGGTCTGCAAAAGAGCAAC 450
QY 61 LeuLeuLeuSerTyHisProProLeuPheArgProMetLysArgIleThrTyrAsnThr 80
DB 451 CTGATTCCTCTCCATCCGCGCTATCTCCGACCATGAGAGCGCATTAACCTGGAAACA 510
QY 81 TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis 100
DB 511 TGGAAAGAGCGCCCTGGTATCCGGGCTCTGGAGAACAGATCGGTATCTCTCTCAT 570
QY 101 ThrAlaTyrAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeuGluVal 120
DB 571 ACAGCTTAATGATGCTGCGCCCAAGGCTCAACACTGCTGTAAGGGCTTGAGAGCT 630
QY 121 CysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGlyAsnHis 140
DB 631 TGTACCTTCAGGCCCATCATCTCTCCAAAGCTCCAAC-TACCCACAGAGAGAAACAC 689
QY 141 ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaValLys 160
DB 690 CGATA-GAATTCACAGCTTAACCTACACCCAGACCTGGAGCAAAAGTCTGTCGANTGAA 748
QY 161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsn 176
DB 749 GGAATTCAGCGAGTTCTGCTCTC--TTTCTGCTAGACTGGTAAT 794
RESULT 13
AK114307 1328 bp mRNA linear INV 30-NOV-2002
LOCUS AK114307
DEFINITION Clona intestinalis cDNA, clone:c1c1054f24, full insert sequence.

```

```

ACCESSION AK114307
VERSION AK114307.1 GI:23585748
KEYWORDS FTI_CDNA.
SOURCE Clona intestinalis
ORGANISM Clona intestinalis
REFERENCE 1
AUTHORS Saitou,Y., Yamada,L., Mochizuki,Y., Takatori,N., Kawashima,T.,
Sasaki,A., Hamaguchi,M., Awazu,S., Yagi,K., Sasaki,Y.,
Nakayama,A., Ishikawa,H., Inaba,K. and Saitoh,N.
TITLE A cDNA resource from the basal chordate Clona intestinalis
JOURNAL Genesis 33 (4), 153-154 (2002)
MEDLINE 22191024
PUBMED 12203911
REFERENCE 2 (bases 1 to 1328)
AUTHORS Saitou,Y. and Saitoh,N.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-2002) Nori Saitoh, Kyoto University, Department of
Zoology, Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
(E-mail:saitoh@ascidian.zool.kyoto-u.ac.jp, Tel:81-75-753-4095,
Fax:81-75-705-1113)
COMMENT Clona intestinalis cDNA Project (URL:
http://ghost.zool.kyoto-u.ac.jp/indexr1.html).
FEATURES
source
1..1328
/organism='Clona intestinalis'
/mol_type='mRNA'
/db_xref='taxon:7719'
/clone='c1c1054f24'
BASE COUNT 481 a 221 c 256 g 370 t
ORIGIN
Alignment Scores:
Pred. No.: 1.37e-60 Length: 1328
Score: 810.00 Matches: 165
Percent Similarity: 64.99% Conservative: 67
Best Local Similarity: 46.22% Mismatches: 103
Query Match: 45.03% Indels: 22
DB: 3 Gaps: 6
US-09-745-506-37 (1-350) x AK114307 (1-1328)
QY 1 MetaspLeuYsAlaLeuLeuSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
DB 82 ATGGAATTGAAGGCTCTCTCTCTGTAATGATGCTTGATCCCTCTGCTGCTGAG 141
QY 21 SerTPAspAsnValGlyLeuLeuValGluProSerProProHisThValAsnThrLeu 40
DB 142 AGCTGGACATGTTGGATTATGTTGTAACCATACCCCAACATACAGTAAGAGATA 201
QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluLysLysAlaAsp 60
DB 202 ATGCTAATCAATGACCTTACTGAGGCTGTCTCAACAGAGGTGACGCCGAAATGTGAC 261
QY 61 LeuLeuLeuSerTyHisProProLeuPheArgProMetLysArgIleThrTyrAsnThr 80
DB 262 ATGATTTGATCTTACCATCCGCTATTTATTTAAGCCTTAAGAGACTAACAATGTCAACA 321
QY 81 TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis 100
DB 322 TGGAAAGAGAGAAATTTGTTGTAATCATTAAGAAACAGATGCAATTTATCCCGCAC 381
QY 101 ThrAlaTyrAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeuGluVal 120
DB 382 ACAGCATGATGATGCTGCTCAAGGTGAGTAATGACTGCTGAGCTTCTGACTGGAGAG 441
QY 121 CysThrSer---ArgProIleHisProSer-----LysAlaProAsnTyrProThr--- 136
DB 442 ATGAATAGCTCAACCAATCATTAAGATCAATTAATTAAGAGATTAATTAATCTGTTCT 501
QY 137 -----GluGlyAsnHisArgValGluPheAsnValAsnTyrThrGlnAspLeuAsp 153
::: |||::: :::

```

```

Db      502 ATCCAAATTCATCAATTAATGACAAAGACTTACTGATTCATTGACAAATCTTAC 561
QY      154 LysValMetSerAlaValysGlyLeaSerValThrPheSerAlaArg 173
Db      562 CAATTTGTAAATTTGTC-----AGC 582
QY      174 ThrGlyAsnGlnThrArgIleAsnLeuAsnCysThrGlnIleValMetGln 193
Db      583 CGTACTAATGTGATGCATAGACATAGATGTAATGCCAATGTGCCAATCTAATAATGAAGAT 642
QY      194 ValValAspPheLeuSerArgAsnLysGlnLeuTyrGlnLysThrGlnIleLeuSerLeu 213
Db      643 ATTTGTAGAT---ATGATGCACACTTAACTGATGATGATGATGATGATGATGATGATGAT 699
QY      214 GlyLysProLeuLeuLeuIleThrGlyMetGlyArgLeuLysThrLeuAspGlnSerVal 233
Db      700 GCTGAACCTCCATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 759
QY      234 SerLeuAlaThrMetIleAspArgIleLysArgIleLysLeuSerHisIleArgLeu 253
Db      760 ACAGTGCAGAACTAATACAGAAATTAATAAACTGCTGCTATTCCTTACGACACTT 819
QY      254 AlaLeuGlyValGlyArgThrLeuGlnSerGlnValLysValAlaLeuLysAlaGly 273
Db      820 TCTCTGGGAGTGTAAACTGTAAACATGATGATGATGATGATGATGATGATGATGATGATGAT 879
QY      274 SerGlySerSerValLeuGlnGlyValGlnAlaAspLeuTyrLeuThrGlyGlnMetSer 293
Db      880 TCTGCTTCTAGCTTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 939
QY      294 HisHisAspThrLeuAspAlaAlaSerGlnGlyIleAsnValIleLeuLysGlnHisSer 313
Db      940 CACCATGAGATATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 999
QY      314 AsnThrGlnArgGlyLeuSerAspLeuAspMetLeuAspSerHisIleu---Glu 332
Db      1000 AACACGAAAGAGGATCTTAAACAGCAAGAAATCTGATGATGATGATGATGATGATGATGAT 1059
QY      333 AsnLysIleAsnIleLeuSerGlnIleThrAspArgAspProLeuGlnVal 349
Db      1060 TCTTTAGTTAAATCATAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1110

RESULT 14
AC037455 170586 bp DNA linear HTG 30-AUG-2001
LOCUS Homo sapiens chromosome 02 clone RP11-422L5, WORKING DRAFT
DEFINITION
AC037455
AC037455.5 GI:9887641
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Smith,D.R.
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 170586)
TITLE Smith,D.R.
REFERENCE Direct Submission
JOURNAL Submitted (09-APR-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
COMMENT On Aug 24, 2000 this sequence version replaced gi:8569072.
-----
Genome Center
Center: Genome Therapeutics Corporation
Center code: GTC
Web site: http://www.genomecorp.com/
Contact: gtc-seqcenter@genomecorp.com
-----
Project Information
Center project name: hg215
-----
Summary Statistics

```

## FEATURES

source

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1171: contig of 1171 bp in length
1172 1271: gap of unknown length
1272 1372: contig of 1048 bp in length
1372 2320 2419: gap of unknown length
2320 3793: contig of 1374 bp in length
3793 3893: gap of unknown length
3893 5414: contig of 1521 bp in length
5414 5515: gap of unknown length
5515 8773: contig of 3259 bp in length
8773 8874: gap of unknown length
8874 10904: contig of 2031 bp in length
10904 10905: gap of unknown length
10905 13204: contig of 2200 bp in length
13204 13305: gap of unknown length
13305 16401: contig of 3097 bp in length
16401 16501: gap of unknown length
16501 18227: contig of 1726 bp in length
18227 18327: gap of unknown length
18327 22401: contig of 4074 bp in length
22401 22502: gap of unknown length
22502 25177: contig of 2676 bp in length
25177 25277: gap of unknown length
25277 29893: contig of 4615 bp in length
29893 29993: gap of unknown length
29993 33457: contig of 3465 bp in length
33457 33557: gap of unknown length
33557 37405: contig of 3848 bp in length
37405 37505: gap of unknown length
37505 41734: contig of 4229 bp in length
41734 41834: gap of unknown length
41834 48999: contig of 7165 bp in length
48999 49000: gap of unknown length
49000 56076: contig of 6977 bp in length
56076 56177: gap of unknown length
56177 63604: contig of 7428 bp in length
63604 70821: gap of unknown length
70821 70921: contig of 7117 bp in length
70921 80138: contig of 9217 bp in length
80138 80238: gap of unknown length
80238 88780: contig of 8542 bp in length
88780 88880: gap of unknown length
88880 104642: contig of 15762 bp in length
104642 104742: gap of unknown length
104742 124656: contig of 19914 bp in length
124656 124757: gap of unknown length
124757 143891: contig of 19135 bp in length
143891 143991: gap of unknown length
143991 170586: contig of 26595 bp in length.
Location/Qualifiers
1. 170586
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="02"
/clone="RP11-422L5"

```

```
misc_feature- /clone_1lb="RPCT-11"
1. .1171
/note="assembly_name:Contig1"
misc_feature 1272. .2319
/note="assembly_name:Contig13"
misc_feature 2420. .3793
/note="assembly_name:Contig17"
misc_feature 3894. .5414
/note="assembly_name:Contig18"
misc_feature 5515. .8773
/note="assembly_name:Contig19"
misc_feature 6874. .10904
/note="assembly_name:Contig20
clone_end:"
misc_feature 11005. .13204
/note="assembly_name:Contig21"
misc_feature 13305. .16401
/note="assembly_name:Contig22"
misc_feature 16502. .18227
/note="assembly_name:Contig23"
misc_feature 18328. .22401
/note="assembly_name:Contig24"
misc_feature 22502. .25177
/note="assembly_name:Contig25"
misc_feature 25278. .23892
/note="assembly_name:Contig26"
misc_feature 29993. .33457
/note="assembly_name:Contig27"
misc_feature 33558. .37405
/note="assembly_name:Contig28"
misc_feature 37506. .41734
/note="assembly_name:Contig29"
misc_feature 41835. .48999
/note="assembly_name:Contig30"
misc_feature 49100. .56076
/note="assembly_name:Contig31
clone_end:sp6"
misc_feature 56177. .63604
/note="assembly_name:Contig32"
misc_feature 63705. .70821
/note="assembly_name:Contig33"
misc_feature 70922. .80138
/note="assembly_name:Contig34"
misc_feature 80239. .88780
/note="assembly_name:Contig35"
misc_feature 88881. .104642
/note="assembly_name:Contig36"
misc_feature 104743. .124656
/note="assembly_name:Contig37"
misc_feature 124757. .143891
/note="assembly_name:Contig38"
misc_feature 143992. .170586
/note="assembly_name:Contig39"
BASE COUNT 50417 a 34173 c 33290 g 50259 t 2447 others
ORIGIN
Alignment Scores:
Pred. No.: 1.04e-51 Length: 170586
Score: 737.00 Matches: 172
Percent Similarity: 37.23% Conservative: 0
Best Local Similarity: 37.23% Mismatches: 1
Query Match: 40.97% Indels: 290
DB: 2 Gaps: 1
US-09-745-506-37 (1-350) x AC037455 (1-170586)
QY 1 MetaspLeuYsAlaLeuLeuSerLeuSnaSpheAlaSerLeuSerPheAlaGlu 20
DB 140069 ATGGATTGGAAGGCTCTCTCTTCCTTGGAATGACTTGGCATCCCTCGTTTCTGAG 140010
QY 21 SerTPaSPaSPaValGlyLeuLeuValGluProSerProPheHisThrValAsnThrLeu 40
DB 140009 AGTTGGGCAATGTTGGATTACTGCTGGACCAAGCCCAACCACTACTGTAAATACACTC 139950
```

```
QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluLysAlaAsp 60
DB 139949 TTCTGACCAATGACTGACTGAGAGAGTGTGAGAGGCTGTCGCAAAAGAGCGAGAC 139890
QY 61 LeuLeuLeuSerTyRHisProProIlePheArgProMetLysArgIleThrPAsnThr 80
DB 139889 CTCATTTCTCTCCATCCATCCGCTATCTCTCGACCAATGAAAGCGCATACCTGGAACACA 139830
QY 81 TrpLysGluArgLeuValIleArgAlaLeuGluSnaTrpValGlyIleTySerProHis 100
DB 139829 TGGAAAGAGCCCTGGTGATGCCGGCTCTGAGAACAGAGTCGCTATCTCCATCAT 139770
QY 101 ThrAlaTyRAspAlaIleProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeu----- 118
DB 139769 ACAGCCTATGATGCTGGGCCCCAGGCGCTCAACAACATGTTGGCTAAAGGGCT-TGGTGA 139711
QY 118 ----- 118
DB 139710 GAAAGCTCTTTCATATTGATATTGATTTTCCCTGCAAAATCTTGAACATTAGAGATGCTTT 139651
QY 118 ----- 118
DB 139650 CCAGTAAAGTTTAACTGCTTTAGGTTGGGAGGCATGATTCATTTTATATATATA 139591
QY 118 ----- 118
DB 139590 TTTTAAATATATGTTGCTGAATTTGGCAGTTAAATGCAATTGMAAAGAAACATTTT 139531
QY 118 ----- 118
DB 139530 TTATTTCTTATGAGTAAGAGAGAGAGAGAGAGAGTCTCTTACTGAGCTTCATGATTTCC 139471
QY 118 ----- 118
DB 139470 TCACTTTGTGGCACAATGCTCTGCACCCAGCAGCACTTGAAATGCTTGTGTAATGAAA 139411
QY 118 ----- 118
DB 139410 GCTCAGGCTCTCCAGANTACATGATGATTAAGTCAAGTAAATAGACATCAGGGTTT 139351
QY 118 ----- 118
DB 139350 AATGAAATGTTTGTAGTAATGTTTACTTATGCTTACAGCAATGAGGGAATATGATGATGGA 139291
QY 118 ----- 118
DB 139290 TTTTAGCTTTATTTTAGTAAGCTTTTACTTATGATTTTGGCTAAAGCTTTTATGATG 139231
QY 118 ----- 118
DB 139230 ATATCTGATTAAGTGTGCAAAATTAAGTATGAGCATTTGCAACCTGATATGATTTAC 139171
QY 118 ----- 118
DB 139170 ATGCTAGATTACCAAGCAGTAGACATTTCTGTCTATTATCTGAAATGATGAAATCAG 139111
QY 118 ----- 118
DB 139110 TCTTTCTTGCAAGTTTTCAGAGATTTAGAAATCTTTGGAATAATATCTGAAATA 139051
QY 118 ----- 118
DB 139050 ATTGATTTTGTATGTCATGCAAAAGCCATAGATTATGAGGAGTGTGATGATATATCTA 138991
QY 118 ----- 118
DB 138990 AGGATGTGTATGATTTATGTAAGCATGAGATGATCTTTTAAAGTCAATTTCTTAG 138931
QY 118 ----- 118
DB 138930 AAATTTGCTATATATAGTTTACTTTTAAATCTTAATCTCTACCCCAAAATCTCCAT 138871
```

```

OY      119      -----GlyAlaCysThrSerArgProIleHisProSerLysAla 131
Db      138870  TTCCCAACAATATTTCTCTAGAGAGCTTACTCCAGGCCCATTAATCTTCCAAAGCT 138811
OY      132      ProbAntyPrProThGluGlyAsnHisArgValGluPheAsnValAsnTyrThrGlnAsp 151
Db      138810  CCCAACTACCCACAGAGGAAACCCACAGTAGAATTCACCTTAACCTAACCCAAAGAC 138751
OY      152      LeuAspLysValMetSerAlaValIleAspGlyValSerPheAlaThrSerPheSer 171
Db      138750  CTGGACAAAGTCATGCTCTCAGTAGAAGAAATGACGGTGTCTCTCCTCTCTTTTCT 138691
OY      172      AlaArg 173
Db      138690  GCTAGG 138685

RESULT 15
LOCUS      AC005037      190508 bp      DNA      linear      PRI 30-SEP-2000
DEFINITION Homo sapiens BAC clone RP11-469M7 from 2, complete sequence.
ACCESSION  AC005037
VERSION     AC005037.2  GI:4827310
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 190508)
            Toward a complete human genome sequence
            Genome Res. 8 (11), 1097-1108 (1998)
            99063792
            9847074
REFERENCE   2 (bases 1 to 190508)
AUTHORS    Abbott, A. and Le,T.
TITLE      The sequence of Homo sapiens BAC clone RP11-469M7
JOURNAL    Unpublished
REFERENCE   3 (bases 1 to 190508)
AUTHORS    Waterston, R.H.
TITLE      Direct Submission
JOURNAL    Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            4 (bases 1 to 190508)
            Waterston, R.H.
            Direct Submission
            Submitted (14-MAY-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            5 (bases 1 to 190508)
            Waterston, R.
            Direct Submission
            Submitted (30-SEP-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On May 14, 1999 this sequence version replaced gi:3109089.
            ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: saplens@wustl.wustl.edu
            ----- Summary Statistics
            -----
            Center project name: H_NH0469M07

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
 The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Fengsen, E., Tatenio, M., Catanesse, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pletier de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)  
 VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:  
 The clone sequenced to the left is RP11-13J8; the clone sequenced to the right is RP11-91M5. Actual start of this clone is at base position 1 of RP11-469M7; actual end is at 190508 of RP11-469M7.

#### FEATURES

source

```

1. 190508
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="2"
   /map="2"
   /clone="RP11-469M7"
   /clone_1kb="RP11-11"
   155..222
   /rpt_family="L2"
   982..1299
   /rpt_family="Alu"
   1502..1533
   /rpt_family="(CA)n"
   1535..1823
   /rpt_family="Alu"
   2005..2308
   /rpt_family="Alu"
   2392..2562
   /rpt_family="L1"
   3284..3315
   /rpt_family="Alu"
   3316..3338
   /rpt_family="(CAA)n"
   3339..3621
   /rpt_family="Alu"
   3650..3743
   /rpt_family="AT-rich"
   4353..4664
   /rpt_family="Alu"
   4727..5018
   /rpt_family="Alu"
   5545..5585
   /rpt_family="(CATA)n"
   5586..5873
   /rpt_family="Alu"
   5876..6007
   /rpt_family="Alu"
   6327..6713
   /rpt_family="MER21_g"
   6924..7294
   /rpt_family="MER2_type"
   8027..8049
   /rpt_family="AT-rich"
   8050..8354
   /rpt_family="Alu"
   8355..8661
   /rpt_family="Alu"

```

repeat\_region 8356.8684  
/rpt\_family="AT-rich"  
repeat\_region 9450.9752  
/rpt\_family="Alu"  
repeat\_region 10371.10680  
/rpt\_family="Alu"  
repeat\_region 10762.11038  
/rpt\_family="Alu"  
repeat\_region 11039.11109  
/rpt\_family="Alu"  
repeat\_region 11138.11410  
/rpt\_family="Alu"  
repeat\_region 11446.11663  
/rpt\_family="L1"  
repeat\_region 11664.11959  
/rpt\_family="Alu"  
repeat\_region 11960.12189  
/rpt\_family="L1"  
repeat\_region 12190.12334  
/rpt\_family="Alu"  
repeat\_region 12335.12360  
/rpt\_family="L1"  
repeat\_region 12464.12514  
/rpt\_family="L1"  
repeat\_region 12515.12822  
/rpt\_family="Alu"  
repeat\_region 12823.12862  
/rpt\_family="L1"  
repeat\_region 12857.13096  
/rpt\_family="L1"  
repeat\_region 13101.13330  
/rpt\_family="MALR"  
repeat\_region 13331.13634  
/rpt\_family="Alu"  
repeat\_region 13635.13692  
/rpt\_family="MALR"  
repeat\_region 13704.13990  
/rpt\_family="L1"  
repeat\_region 14082.14387  
/rpt\_family="Alu"  
repeat\_region 14402.14704  
/rpt\_family="Alu"  
repeat\_region 14797.15095  
/rpt\_family="Alu"  
repeat\_region 15162.15196  
/rpt\_family="Mariner"  
repeat\_region 15606.15681  
/rpt\_family="L1"  
repeat\_region 15691.15827  
/rpt\_family="MERL\_type"  
repeat\_region 15861.16189  
/rpt\_family="L2"  
repeat\_region 16426.16446  
/rpt\_family="(CAAT)n"  
repeat\_region 16447.16730  
/rpt\_family="Alu"  
repeat\_region 17072.17179  
/rpt\_family="L2"  
repeat\_region 17212.17504  
/rpt\_family="Alu"  
repeat\_region 17869.18139  
/rpt\_family="Alu"  
repeat\_region 18168.18464  
/rpt\_family="Alu"  
repeat\_region 18606.19155  
/rpt\_family="MALR"  
repeat\_region 19454.19767  
/rpt\_family="Alu"  
repeat\_region 19807.19949  
/rpt\_family="MERL\_type"  
repeat\_region 19950.20247  
/rpt\_family="Alu"

Alignment Scores:  
Pred. NO.: 1.19e-51 Length: 190508  
Score: 737.00 Matches: 172  
Percent Similarity: 37.23% Conservative: 0  
Best local Similarity: 37.23% Mismatches: 1  
Query Match: 40.97% Indels: 290  
DB: 9 Gaps: 1  
US-09-745-506-37 (1-350) x AC005037 (1-190508)  
OY 1 MetaspLeuysaLeuLeuSerLeuaspPhealSerLeuSerPhealGlu 20  
DB 122461 ATGATTTGAAGCTCTCCTTCTTCTTGATGACTTGATCCCTGCTGTCGAG 122402  
OY 21 SerTpsaspasValGlyLeuLeuValGluProSerPropOHisthrValasThrLeu 40  
DB 122401 AGTTGGACATGTTGATTCGTGTGGAACCAAGCCACACATCTGTAATAATCACTC 122342  
OY 41 PheLeuThrPheaspLeuThrGluGluValMetGluGluValLeuGlnLysAlaASP 60  
DB 122341 TTCTGACCAATGACCTGACTGAGGAAGTGTGAGAGGTGCTGCCAAGGAAGCAGAC 122282  
OY 61 LeuLeuSerTyrHisProPioLePheArgProLeuLysArgLLeuThrPasnThr 80  
DB 122281 CTCATTCCTCTACACATCCGCTATCTCCGACCAATGAAGCCATTAACCTGGAACACA 122222  
OY 81 TrpLysGluArgLeuValLeuArgAlaLeuGlnasnArgValGlyLeuSerProHis 100  
DB 122221 TGGAGGAGCCCGGTGATCCGGGCTGTGGAGAACAGTGGTATCTCTCTCAT 122162  
OY 101 ThrAlaTyrAspAlaAlaProGlnGlyValasnThrPhealLysGlyLeu----- 118  
DB 122161 ACAGCCTATGATGCTGCCGCCAGGCGCTCAACAACGTGTGGCTAAAGGCT-TGGTGA 122103  
OY 118 ----- 118  
DB 122102 GAAGCCTTCATATTTGATATTTTCCCTCAATACTTTGAAACTTAGATGCTTT 122043  
OY 118 ----- 118  
DB 122042 CCAGTAAGTTTAACTGCTTTAGGGTTGGGAGGCCATGATTCATTTTATATATA 121983  
OY 118 ----- 118  
DB 121982 TTTTAATATATGTTGCTGAATTTGGCAGTTAATTCATTTGAAAAAGAACATTTT 121923  
OY 118 ----- 118  
DB 121922 TTATTCCTATGAGTAGAGAGAGACAGAGTCTTCCTACTGAGCTGACCTCAGATTC 121863  
OY 118 ----- 118  
DB 121862 TCACCTTTGTGGCACAATGCCCTGTCACCCAGGCACTTGGAAAAATGCTTGTGAATGAA 121803  
OY 118 ----- 118  
DB 121802 GCTCAGGCGCTCCAGAAATACACTGATTTAGTAACTCAAGTAATAGACATCAGGTTT 121743  
OY 118 ----- 118  
DB 121742 AATGATGTTTTAGTGAATGTTTACTTATGCCTAGACAAATGAGGAATGATGATGA 121683  
OY 118 ----- 118  
DB 121682 TTTTACCTTATTTAGTATGATCTTTAGTTTACATTTTGGGCTTAAAGCTTTTATGATG 121623  
OY 118 ----- 118  
DB 121622 AATCTGATTAATGTCGCAATAAAGTATGACATTTGCAACTGATGATGTTTAC 121563  
OY 118 ----- 118  
DB 121562 ATGCTAGATTACCAAGCAGTAGACCATTTCTGTATATCTGAAAAATGTATGATCAG 121503



```
QY 118 ----- 118
Db 121502 TCTTCTGTCCAGTTTTCAGAGATTAGAAATCTTTGGAAAAATAATCTGAAATA 121443
QY 118 ----- 118
Db 121442 ATTTGATTTGTATGTCATGCCAAAGCCCTAGATTANGGGAGCTGGTGTATATCTTA 121383
QY 118 ----- 118
Db 121382 AGGGATGTGTATGTATGTATAGCATGCAATGAAATCTTTTAAAAAGTCATTCTTAG 121323
QY 118 ----- 118
Db 121322 AAATTTGCCATATAATAGTTTACTTTTAAATCTTAATCTCTCACCCCCCAAACTCCAT 121263
QY 119 ----- GlyAlacysThrSerArgProIleHisPrcSerLysAla 131
Db 121262 TTTCCACAAATATTTCTCTAGAGCTGTGTACCTCCAGGCCCATCATCTCCAAAGCT 121203
QY 132 ProAsnTyrProThrGluGlyAsnHisArgValGluPheAsnValAsnTyrThrGlnAsp 151
Db 121202 CCCAACTACCTCTACAGAGGAAACACCGAGTAGAATTCACGTTTACTACACCCAGAGAC 121143
QY 152 LeuAspLysValMetSerAlaValIysGlyIleAspGlyValSerValThrSerPheSer 171
Db 121142 CTGGACAAAGTCATGTCTGCACATGAAGAATGACGCGTGTCTGTCTCATCTTTTCT 121083
QY 172 AlaArg 173
Db 121082 GCTAGG 121077
```

Search completed: August 23, 2003, 19:24:38  
Job time : 3070 secs

**THIS PAGE BLANK (USPTO)**